

Sampling error distribution for the ensemble Kalman filter update step

Andrey Kovalenko · Trond Mannseth · Geir Nævdal

Received: 31 January 2011 / Accepted: 30 May 2011 / Published online: 6 July 2011
© The Author(s) 2011. This article is published with open access at Springerlink.com

Abstract In recent years, data assimilation techniques have been applied to an increasingly wider specter of problems. Monte Carlo variants of the Kalman filter, in particular, the ensemble Kalman filter (EnKF), have gained significant popularity. EnKF is used for a wide variety of applications, among them for updating reservoir simulation models. EnKF is a Monte Carlo method, and its reliability depends on the actual size of the sample. In applications, a moderately sized sample (40–100 members) is used for computational convenience. Problems due to the resulting Monte Carlo effects require a more thorough analysis of the EnKF. Earlier we presented a method for the assessment of the error emerging at the EnKF update step (Kovalenko et al., SIAM J Matrix Anal Appl, in press). A particular energy norm of the EnKF error after a single update step was studied. The energy norm used to assess the error is hard to interpret. In this paper, we derive the distribution of the Euclidean norm of the sampling error under the same assumptions as before, namely normality of the forecast distribution

and negligibility of the observation error. The distribution depends on the ensemble size, the number and spatial arrangement of the observations, and the prior covariance. The distribution is used to study the error propagation in a single update step on several synthetic examples. The examples illustrate the changes in reliability of the EnKF, when the parameters governing the error distribution vary.

1 Introduction

Maximizing the hydrocarbon production requires substantial knowledge of the reservoir. Properties of reservoir rocks and fluids are used in mathematical models, which are in turn used to predict the reservoir's future behavior. Accurate predictions are essential for successful hydrocarbon recovery. Since direct observation of each and every point in the reservoir is technically impossible, one cannot populate a reservoir model with exactly known values. The observation techniques used to acquire the reservoir properties before production starts are either local (wireline logs and formation tests) or to a large degree imprecise (seismic surveys). Due to these reasons, reservoir models operate with uncertain parameters and forecasts of the reservoir behavior are, in turn, also uncertain. Reducing the uncertainty of the forecasts plays a crucial role in the hydrocarbon recovery process.

In order to improve the forecasts, the reservoir model needs to be adjusted with the information from production data. Time-lapse seismic data can also be used in updating reservoir models, see, e.g., [1, 6, 13]. Adjusting the model parameters based on the observed

A. Kovalenko (✉) · T. Mannseth
Uni CIPR, P.O. Box 7810, 2050 Bergen, Norway
e-mail: andrey.kovalenko@uni.no

T. Mannseth
e-mail: trond.mannseth@uni.no

A. Kovalenko · T. Mannseth · G. Nævdal
Department of Mathematics, University of Bergen,
Bergen, Norway

G. Nævdal
International Research Institute of Stavanger,
Bergen, Norway
e-mail: Geir.Naevdal@iris.no

data is known as history matching. Improving the uncertain reservoir model is done by assuming the reservoir fluid and rock properties to be random values. The reservoir states, i.e., pressures and saturations, are in this case also random. A widely used methodology to condition the states and parameters on observations is the Bayesian framework. The poorly known reservoir states and parameters are associated with their respective probability density functions. Combining the system state, $u(x, t)$, and the system parameters, $\mu(x)$, into a single vector $\theta(x, t) = (u(x, t), \mu(x))^T$, one can define a joint density function, $f(\theta)$. Further, for the observations, denoted as $d(x, t)$, one defines a likelihood distribution $f(d|\theta)$. Using Bayes' rule, one may directly write

$$f(\theta|d) \propto f(\theta)f(d|\theta). \quad (1)$$

Since history matching is based on sequential data sets (production records, time-lapse seismic surveys, etc.), it is natural to adjust the reservoir model as the data are available. The sequential process of updating the system state and parameters by incorporating a series of measured data into the numerical model is called data assimilation. The purpose of data assimilation in the Bayesian setting is to find the posterior joint probability distribution of the reservoir states and parameters conditioned on the series of observed dynamic data.

Generally, the extended model state, $\theta(x, t)$, is a continuous function of time. It is convenient to work with a model state discretized in time, i.e., $\theta(x, t)$ is represented at certain time intervals as $\theta_k(x) = \theta(x, t_k)$, where $k = 0, \dots, N_t$, and N_t is the total number of time instances. The extended model state is also discretized in space such that $\theta_k = (\theta_k(x_1), \dots, \theta_k(x_M))^T$, where M is the spatial model dimension. The data are represented such that $d_k = d_k(x_1), \dots, d_k(x_{N_d})$, where N_d is the dimension of the resulting data vector. For the sake of simplicity, we will further assume the extended model state to be discretized in time such that the data arrive at the same instances, t_k . Following Evensen and van Leeuwen [11], we write the expression for the conditional probability function for the discrete time case:

$$f(\theta_k|d_k, \dots, d_1) \propto f(\theta_k|d_{k-1}, \dots, d_1)f(d_k|\theta_k). \quad (2)$$

Reducing Eq. 1 to the discrete form (Eq. 2) is done under the assumption of the model evolution to be a first order Markov process, i.e., θ_k depends only on θ_{k-1} , but not on any of the preceding states, $\theta_{k-2}, \dots, \theta_0$.

In general, finding the posterior density in Eq. 2 is difficult. In order to make it computationally feasible,

it is assumed that the prior distribution of the extended state vector is Gaussian. The density $f(\theta_k|d_{k-1}, \dots, d_1)$ is therefore assumed to be Gaussian with mean θ_k^f and covariance C_k^f , denoted as $\mathcal{N}(\theta_k^f, C_k^f)$. Assuming a linear relationship between model parameters and observations through the measurement matrix, H , and adding independent Gaussian noise, i.e.,

$$d_k = H\theta_k + \varepsilon_k^d, \quad \varepsilon_k^d \sim \mathcal{N}(0, C_k^d),$$

is sufficient for the posterior distribution to be Gaussian. The posterior mean and covariance maximize the likelihood function given in Eq. 2. Omitting the details which can be found, for example, in [10], we may write the equations for the posterior mean, θ_k^a , and posterior covariance, C_k^a :

$$\theta_k^a = \theta_k^f + K_k(d_k - H\theta_k^f), \quad (3)$$

$$C_k^a = (I - K_k H)C_k^f. \quad (4)$$

$$K_k = C_k^f H^T (H C_k^f H^T + C_k^d)^{-1}, \quad (5)$$

Here, K is the Kalman gain matrix. Equations 3–5 are known as the Kalman filter analysis equations [19]. These equations are used to obtain the posterior distribution of the extended state vector taking into account the observations at the k th time increment. In order to propagate the system forward in time, the forecast equations are applied:

$$\theta_k^f = A_k \theta_{k-1}^a + \varepsilon_k^m, \quad \varepsilon_k^m \sim \mathcal{N}(0, C_k^m), \quad (6)$$

$$C_k^f = A_k C_{k-1}^a A_k^T + C_k^m, \quad (7)$$

where A is a linear operator of system evolution and ε^m is a normally distributed model error.

For a Gaussian extended state vector and a linear operator of system evolution, the Kalman filter delivers the best estimate. In a great number of real problems, application of the Kalman filter is computationally unfeasible due to the size of the system, even if the forward model is linear. The covariances are too large, and the corresponding matrix equations are difficult to solve correctly. In case of a non-linear forward model, the assumption of Gaussianity that the Kalman filter relies on is not satisfied. Applying such an operator to the Gaussian system state results in a non-Gaussian output, thus making direct assessment of the resulting distribution impossible.

In order to deal with these two issues, a Monte Carlo technique is applied. Instead of adjusting the distribution by running the Kalman filter update and forecast Eqs. 3–7, one estimates the mean and covariance from a moderately sized ensemble. Even though little is known about the ability of such a technique to fully describe non-Gaussian distributions, ensemble-based methods are applied to problems with non-linear forward operator.

The most commonly applied data assimilation method is the ensemble Kalman filter (EnKF), first introduced by Evensen [7]. The method is easy to apply to different state and parameter estimation problems. An extensive description can be found in [10]. It has been applied, for example, to problems in climatology, oceanography, and meteorology, see, e.g., [3, 8, 18].

In oil recovery, the EnKF has been applied for reservoir characterization and history matching, see, e.g., [22, 23, 26]. For an extensive literature review and discussion of the methodology, see [2]. For a review of recent applications, see [25]. It has also been used within closed-loop reservoir management and production optimization, see, e.g., [4] and [24].

The EnKF starts out with an ensemble $\{\theta_{i,0}^a\}_{i=1}^{N_e}$ from the initial distribution. For a (generally nonlinear) model operator, F_k , each ensemble member is updated as follows:

$$\theta_{i,k}^f = F_k(\theta_{i,k-1}^a) + \varepsilon_{i,k}^m, \quad (8)$$

where $\varepsilon_{i,k}^m$ are the modeling errors, generated independently for each sample. The covariance is estimated from the resulting ensemble:

$$\tilde{C}_k^f = \frac{1}{N_e - 1} \sum_{i=1}^{N_e} (\theta_{i,k}^f - \bar{\theta}_k^f) (\theta_{i,k}^f - \bar{\theta}_k^f)^T, \quad (9)$$

where $\bar{\theta}$ denotes the ensemble mean. Replacing the true covariance with its estimate and adding independent errors to the observed data (the resulting vectors are denoted as $d_{i,k}$), one performs the analysis stage as in Eqs. 3 and 5:

$$\theta_{i,k}^a = \theta_{i,k}^f + \tilde{K}_k (d_{i,k} - H\theta_{i,k}^f), \quad (10)$$

$$\tilde{K}_k = \tilde{C}_k^f H^T (H \tilde{C}_k^f H^T + C_k^d)^{-1}. \quad (11)$$

The analyzed sample mean, $\bar{\theta}_k^a$, and the corresponding sample covariance, \tilde{C}_k^a , computed from the analyzed ensemble as in Eq. 9 approximate the posterior mean and covariance. It is worth pointing out that in general even the exact mean and covariance do not fully characterize non-Gaussian distributions, but this issue is not

considered in the current article. Instead, we focus on the Monte Carlo approximation of the analysis distribution. Further in the article, we will limit ourselves to the case of linear forward and measurement operators.

Ensemble-based estimation leads to several issues. Sampling error affects both mean and covariance estimates but vanishes as the ensemble size tends to infinity. With EnKF, the sampling error depends not only on the ensemble size but also on the number and mutual locations of measurements, see, for example, [12].

Sampling errors in the covariance matrix estimate lead to spurious correlations, i.e., erratic changes of the state vector components in regions of no real influence. The most common approach to avoid spurious correlations is to perform the update locally, i.e., the extended state vector components are updated only if they are close to the observations [14, 18]. In practice, closeness is often defined somewhat “ad hoc”. If the number of observations exceeds the ensemble size, the entire ensemble can collapse, i.e., the spread of the ensemble decreases gradually with every time step. Ensemble collapse and localization techniques, however, are not addressed in the current paper.

With the finite ensemble size, EnKF may experience rank deficiency. The issue arises due to the lack of the degrees of freedom (number of ensemble members) to properly assimilate the measurements, if the ensemble size is smaller than the number of measurements. A thorough discussion on how to apply the EnKF or its variants in this situation can be found in [9, 10]. We will consider the case when the ensemble size is greater than the number of measurements further in the article.

The EnKF is also affected by the relation between the size of the ensemble and the number of assimilated data points. In [20], e.g., it has been shown that increasing the number of measurements leads to degeneracy of the ensemble. In the numerical studies performed in [12], it has been shown that increasing number of measurements leads to loss of quality of the EnKF update.

The ensemble of states and parameters updated with the EnKF provides an estimate of the uncertainty associated with the reservoir. However, very little is known about the accuracy of the method itself. It is possible to build an empirical distribution function for the EnKF sampling error, but this approach would require a huge computational effort. Thus, in order to approximate the cumulative distribution function with N_s points, one needs to run $N_e \times N_s$ EnKF updates. In addition, such an approximation would also be extremely problem dependent.

Better understanding of the quality of the EnKF is the purpose of the present paper. We focus on an essential part of the EnKF algorithm, namely the analysis step. We study the norm of the sampling error, defined as the difference between the outcomes from the Kalman filter analysis equations and their EnKF counterparts.

In Kovalenko et al., submitted, we obtained a distribution for a particular energy norm of the sampling error arising at a single EnKF analysis step. We used several results from random matrix theory in order to derive the distribution. The norm of the sampling error we used in Kovalenko et al., submitted, however, is not easy to interpret. In [21], we gave a sketch of the derivation of the distribution for the Euclidean norm.

In the current paper, we derive the distribution of the Euclidean norm in full details. The distribution depends on the system size, number of ensemble members, prior covariance properties, and mutual locations of the measurements. In order to obtain the distribution, we make the same assumptions as in Kovalenko et al., submitted. Namely, we assume the measurement errors to be negligible, the forecast distribution to be Gaussian, and the model and measurement operators to be linear. Although one cannot expect to know the exact forecast distribution in a real reservoir characterization problem, this simplistic setting is still of interest as it can be considered as a “best-case scenario” for the EnKF. It is, however, of great significance to extend the obtained distribution to a real reservoir case.

We also perform a set of numerical experiments illustrating the behavior of the obtained distribution. For a given reference field and a given prior distribution, we build the distribution of the sampling error in various situations, as if we were to perform the EnKF updates of the state vector. We vary the parameters the distribution depends on in order to investigate their influence on the quality of the EnKF update. The analytical distribution is then compared with the empirical distribution of the sampling error norm.

The paper is organized as follows: Section 2 gives an overview of certain results from random matrix theory used to derive the main result. Section 3 contains the derivation of the main result. The section is split into two parts. In the first subsection, we give a shortened derivation of the result from Kovalenko et al., submitted, given here for the sake of completeness. The second subsection contains the extension of the derivation to the case of the Euclidean norm, given in full details. Section 4 is dedicated to the numerical experiments. Finally, we give an outlook and review the possibilities for future extensions.

2 Some properties of random matrices

The EnKF analysis step, given by Eqs. 10–11, is based on the sample covariance matrix, given in Eq. 9. Due to the crucial role of the sample covariance in the analysis step, it is worth noting several properties of this matrix. The results below along with the corresponding proofs can be found in [15, 27].

Omitting the time index from now on, assume the ensemble members, $\{\theta_i^f\}_{i=1}^{N_e}$, to be independent random samples from the Gaussian distribution, $\mathcal{N}(\theta^f, C)$. Consider then the following matrix, $\Theta = (\theta_1^f, \dots, \theta_{N_e}^f)$. The resulting random matrix is said to have a matrix variate normal distribution and is denoted as follows:

$$\Theta \sim \mathcal{N}_{M, N_e}(\theta^f e^T, C \otimes I_{N_e}), \quad (12)$$

where $e^T(1 \times N_e) = (1, \dots, 1)$ and $C(M \times M) > 0$ is the prior covariance. The symbol \otimes denotes Kronecker matrix product, i.e.,

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \dots & a_{1n}B \\ a_{21}B & a_{22}B & \dots & a_{2n}B \\ \vdots & & & \\ a_{m1}B & a_{m2}B & \dots & a_{mn}B \end{pmatrix},$$

where $A(m \times n) = (a_{ij})$ and $B(p \times q) = (b_{ij})$.

The matrix generalization of the normal distribution preserves some useful properties of the vector multivariate normal distribution. In particular, the matrix multivariate normal distribution is invariant with respect to certain linear transforms. If a matrix Y is normally distributed as $Y \sim \mathcal{N}_{p,n}(M, \Sigma \otimes \Psi)$ and is multiplied by a deterministic matrix $D(m \times p)$ of rank $m \leq p$ from the left and by another deterministic matrix $G(n \times t)$ of rank $t \leq n$ from the right, then the resulting product also has matrix multivariate normal distribution,

$$DYG \sim \mathcal{N}_{m,t}(DMG, (D\Sigma D^T) \otimes (G^T\Psi G)). \quad (13)$$

The transpose of a normally distributed matrix $Y \sim \mathcal{N}_{p,n}(M, \Sigma \otimes \Psi)$ is again normally distributed as follows:

$$Y^T \sim \mathcal{N}_{n,p}(M^T, \Psi \otimes \Sigma). \quad (14)$$

Denote $\bar{\theta}^f$ to be the ensemble sample mean and consider the matrix $X = \Theta - \bar{\theta}^f e^T = (\theta_1 - \bar{\theta}^f, \dots, \theta_{N_e} - \bar{\theta}^f)$. The matrix product, $\frac{1}{N_e-1}XX^T$, is exactly the sample covariance matrix given in Eq. 9, and we write

$$\tilde{C} = \frac{1}{N_e-1}XX^T. \quad (15)$$

It can be shown that X is distributed as $\mathcal{N}_{M, N'_e}(0, C \otimes I_{N'_e})$, where $N'_e = N_e - 1$, and is independent of $\bar{\theta}^f$.

The product of a normally distributed matrix, $Y \sim \mathcal{N}_{p,n}(0, \Sigma \otimes I)$, and its transpose is said to have Wishart distribution and is denoted as $W_p(n, \Sigma)$, where p is the dimension of the matrix and n is the number of degrees of freedom. The sample covariance matrix, \tilde{C} , satisfies this definition due to Eq. 15 and belongs to Wishart distribution $W_M(N'_e, \frac{C}{N'_e})$. If the number of degrees of freedom is less than the dimension of the Wishart matrix, the distribution is said to be singular, and such random matrices are not invertible, but some useful properties of the regular Wishart distribution are still retained [27].

3 Distribution of the sampling error

The EnKF approximates the joint distribution of the states and parameters from a moderately sized ensemble. Due to approximation of the covariance from the ensemble, an error inevitably occurs. While the sampling error at the forecast step is easy to estimate, the error at the analysis step is more complex to assess.

In the current article, we focus on the error the EnKF itself produces at the analysis step. The first assumption we make is to consider the entire ensemble to be independent realizations from the correct forecast distribution, $\mathcal{N}(\theta^f, C)$, $\theta^f \in \mathbb{R}^M$, $C(M \times M) > 0$.

In order to focus on the error arising purely at the analysis step, we assume that the ensemble mean is equal to the true expectation, i.e., $\bar{\theta}^f = \theta^f$. For the sake of convenience, we enumerate the state vector so that the first N_d entries are the observed components. Thus, the first N_d columns of the $N_d \times M$ matrix H form an identity matrix, and the remaining $M - N_d$ columns consist of zeros. The covariance matrix is split as follows:

$$C = \begin{pmatrix} C_H & C_1^T \\ C_1 & C_2 \end{pmatrix}, \quad (16)$$

where $C_H = HCH^T$. The corresponding sample covariance matrix is assumed to be split accordingly. We consider negligible measurement errors, thus letting $C_d = 0$. This also implies $d_i = d$. Finally, we consider $N_e > N_d + 1$.

Under these assumptions, we consider the outcomes from the Kalman filter and the EnKF analysis equations, θ^a and $\bar{\theta}^a$, correspondingly. Since the Kalman filter provides the correct posterior distribution in this simplistic setting, the sampling error is the difference between the outcomes from the EnKF and the Kalman

filter. Denoting $\delta = d - H\bar{\theta}^f$, we write the expression for the sampling error:

$$\bar{\theta}^a - \theta^a = (\tilde{K} - K)\delta. \quad (17)$$

Below we will derive the distribution of the norm of the sampling error given in Eq. 17, using the results from random matrix theory discussed in the previous section.

The first step would be to find the matrix distribution of the difference between the Kalman gain and the ensemble Kalman gain. Splitting the true covariance and its ensemble estimate as in Eq. 16, we may write:

$$\begin{aligned} \|\tilde{K} - K\| &= \left\| \begin{pmatrix} \tilde{C}_H \\ \tilde{C}_1 \end{pmatrix} \tilde{C}_H^{-1} - \begin{pmatrix} C_H \\ C_1 \end{pmatrix} C_H^{-1} \right\| \\ &= \|\tilde{C}_1 \tilde{C}_H^{-1} - C_1 C_H^{-1}\|. \end{aligned}$$

The matrix \tilde{C}_H has Wishart distribution $W_{N_d}(N'_e, \frac{C_H}{N'_e})$. It is nonsingular if and only if $N_e > N_d + 1$, and from now on we limit ourselves to this case. According to [15, 27], the matrix \tilde{C}_1 has the following marginal distribution:

$$\tilde{C}_1 | \tilde{C}_H \sim \mathcal{N}_{M-N_d, N_d} \left(C_1 C_H^{-1} \tilde{C}_H, C_{2,2} \otimes \frac{\tilde{C}_H}{N'_e} \right),$$

where $C_{2,2} = C_2 - C_1 C_H^{-1} C_1^T$. Multiplying \tilde{C}_1 by \tilde{C}_H^{-1} from the right, we use Eq. 13 with $G = \tilde{C}_H^{-1}$ and subtract the corresponding mean:

$$(\tilde{K} - K) | \tilde{C}_H \sim \mathcal{N}_{M-N_d, N_d} \left(0, C_{2,2} \otimes \frac{\tilde{C}_H^{-1}}{N'_e} \right). \quad (18)$$

Consider the distribution of $Z = (\tilde{K} - K)^T C_{2,2}^{-1/2}$. According to Eqs. 14 and 18,

$$(\tilde{K} - K)^T | \tilde{C}_H \sim \mathcal{N}_{M-N_d, N_d} \left(0, \frac{\tilde{C}_H^{-1}}{N'_e} \otimes C_{2,2} \right).$$

By Eq. 13, substituting $Y = (\tilde{K} - K)^T$ and $G = C_{2,2}^{-1/2}$,

$$Z | \tilde{C}_H \sim \mathcal{N}_{M-N_d, N_d} \left(0, \frac{\tilde{C}_H^{-1}}{N'_e} \otimes I_{M-N_d} \right). \quad (19)$$

In the following subsection, we give the derivation for a special energy norm of the sampling error. The case of the Euclidean norm is considered in Section 3.2.

3.1 Distribution of the energy norm

Consider a norm $\|\cdot\|_{C_{2,2}}$ such that for an arbitrary vector $u \in \mathbb{R}^{M-N_d}$, $\|u\|_{C_{2,2}}^2 = (C_{2,2}^{-1}u, u)$. One may then write:

$$\begin{aligned} \|(\tilde{K} - K)\delta\|_{C_{2,2}}^2 &= \delta^T (\tilde{K} - K)^T C_{2,2}^{-1} (\tilde{K} - K) \delta = (Z Z^T \delta, \delta). \end{aligned} \quad (20)$$

With the same argument as in Eq. 15, the matrix product, ZZ^T , has Wishart distribution, namely

$$ZZ^T|\tilde{C}_H \sim W_{N_d}\left(M - N_d, \frac{\tilde{C}_H^{-1}}{N'_e}\right). \quad (21)$$

In order to find the distribution of the quadratic form given in Eq. 20, one can exploit another useful property of Wishart matrices, namely the relation between the Wishart distribution and the χ^2 distribution [15]. For any Wishart matrix $S \sim W_p(n, \Sigma)$ and any random vector $y(p \times 1)$ distributed independently of S , such that $P(y \neq 0) = 1$, the following ratio has χ^2 distribution:

$$\frac{(Sy, y)}{(\Sigma y, y)} \sim \chi_n^2 \quad (22)$$

and is independent of y . A similar result holds for the inner product ratio for the inverted Wishart matrix:

$$\frac{(\Sigma^{-1}y, y)}{(S^{-1}y, y)} \sim \chi_{n-p+1}^2. \quad (23)$$

By Eqs. 22 and 23, for an arbitrary non-zero $u \in \mathbb{R}^{N_d}$, independent of ZZ^T and \tilde{C}_H , the following holds:

$$\frac{(N'_e ZZ^T u, u)}{(\tilde{C}_H^{-1} u, u)}|\tilde{C}_H \sim \chi_{M-N_d}^2, \quad (24)$$

$$\frac{(N'_e C_H^{-1} u, u)}{(\tilde{C}_H^{-1} u, u)} \sim \chi_{N_e-N_d+1}^2. \quad (25)$$

Since the marginal distribution of the matrix ZZ^T is Wishart (given \tilde{C}_H), we first consider the conditional probability $P((ZZ^T u, u) < \varepsilon|\tilde{C}_H)$. Multiplying both sides of the inequality under the probability sign by certain values in order to obtain the ratios of quadratic forms as above, we write:

$$\begin{aligned} P\left(\frac{(N'_e ZZ^T u, u)}{(\tilde{C}_H^{-1} u, u)} < \frac{\varepsilon}{(C_H^{-1} u, u)} \frac{(N'_e C_H^{-1} u, u)}{(\tilde{C}_H^{-1} u, u)}|\tilde{C}_H\right) \\ = P\left(\xi_{\chi_{M-N_d}^2} < \varepsilon' \frac{(N'_e C_H^{-1} u, u)}{(\tilde{C}_H^{-1} u, u)}|\tilde{C}_H\right) \\ = \int_0^{\varepsilon' \frac{(N'_e C_H^{-1} u, u)}{(\tilde{C}_H^{-1} u, u)}} P_{\chi_{M-N_d}^2}(y) dy, \end{aligned}$$

where $\varepsilon' = \frac{\varepsilon}{(C_H^{-1} u, u)}$. The unconditional probability is then an expectation over all possible values of the Wishart matrix \tilde{C}_H :

$$\begin{aligned} P((ZZ^T u, u) < \varepsilon) \\ = \int_{\Lambda > 0} p_{W_{N_d}}(N'_e, \frac{C_H}{N'_e})(\Lambda) \int_0^{\varepsilon' \frac{(N'_e C_H^{-1} u, u)}{(\Lambda^{-1} u, u)}} p_{\chi_{M-N_d}^2}(y) dy d\Lambda, \end{aligned} \quad (26)$$

where $p_{W_{N_d}}(N'_e, C_H)(\Lambda)$ is the density function of the Wishart distribution.

It can be shown that

$$\begin{aligned} \int_{\Lambda > 0} p_{W_{N_d}}(N'_e, \frac{C_H}{N'_e})(\Lambda) \int_0^{\varepsilon' \frac{(N'_e C_H^{-1} u, u)}{(\Lambda^{-1} u, u)}} p_{\chi_{M-N_d}^2}(y) dy d\Lambda \\ = \int_0^\infty p_{\chi_{N_e-N_d+1}^2}(z) \int_0^{\frac{z\varepsilon}{(C_H^{-1} u, u)}} p_{\chi_{M-N_d}^2}(y) dy dz. \end{aligned} \quad (27)$$

Equality Eq. 27 is a technical result and could be proved by using Eq. 25. Substituting $u = \delta$ in Eq. 27 and using Eqs. 20 and 17, we formulate

Theorem 1 Let $\bar{\theta}^a$ and θ^a be the outcomes from the analysis step of the EnKF and the Kalman filter, respectively. Let $\delta = d - H\bar{\theta}^f$. Define $\|u\|_{C_{2,2}} = \|C_{2,2}^{-1/2}u\|$. Then

$$\begin{aligned} P(\|\bar{\theta}^a - \theta^a\|_{C_{2,2}}^2 < \varepsilon) \\ = \int_0^\infty p_{\chi_{N_e-N_d+1}^2}(z) \int_0^{\frac{z\varepsilon}{(C_H^{-1} \delta, \delta)}} p_{\chi_{M-N_d}^2}(y) dy dz. \end{aligned} \quad (28)$$

The detailed proof of Theorem 1 is available in Kovalenko et al., submitted. Using Eq. 28, the probability of the EnKF outcome to deviate from the reference can be computed. However, the closeness of the EnKF outcome to the reference is defined by the norm $\|\cdot\|_{C_{2,2}}$, which is hard to interpret.

3.2 Distribution of the Euclidean norm

Below we will find the distribution of the Euclidean norm of the sampling error, $\bar{\theta}^a - \theta^a$. Using the definition of Z , we write:

$$\|\bar{\theta}^a - \theta^a\|^2 = (Z C_{2,2} Z^T \delta, \delta). \quad (29)$$

In order to derive the distribution for the latter norm, we consider the following decomposition of the matrix $C_{2,2}$:

$$C_{2,2} = \sigma_1^2 P_1 + \cdots + \sigma_{M-N_d}^2 P_{M-N_d}, \quad (30)$$

where σ_i is a singular value of $C_{2,2}$ and P_i is a corresponding projection matrix [16]. Such a decomposition is called a spectral decomposition of a symmetric matrix, and the decomposition is unique. For the sake of simplicity, we will consider the case when the matrix $C_{2,2}$ has unequal singular values, and the corresponding projection matrices are all of rank 1. A similar result for the general case is easily obtained by following the same proof with minor adjustments.

Applying the spectral decomposition to the matrix in the inner product in Eq. 29, we get:

$$\begin{aligned} (Z C_{2,2} Z^T \delta, \delta) &= \sum_{i=1}^{M-N_d} \sigma_i^2 (Z P_i Z^T \delta, \delta) = \sum_{i=1}^{M-N_d} \sigma_i^2 (Z P_i^2 Z^T \delta, \delta) = \\ &= \sum_{i=1}^{M-N_d} \sigma_i^2 (P_i Z^T \delta, P_i Z^T \delta). \end{aligned} \quad (31)$$

Since $P_i P_j = 0, i \neq j$, the sum terms in Eq. 31 are independent, and we will consider the distribution of a single inner product, $(P_i Z^T \delta, P_i Z^T \delta)$, thus finding the distribution for the norm of the sampling error. Since $Z|\tilde{C}_H$ is a normally distributed matrix, we use Eq. 13 setting $D = \delta^T$ and write the following marginal distribution for the random vector $P_i Z^T \delta$:

$$P_i Z^T \delta | \tilde{C}_H \sim \mathcal{N} \left(0, \frac{(\tilde{C}_H^{-1} \delta, \delta)}{N'_e} \times P_i^2 \right).$$

Normalizing the variance, we obtain:

$$\frac{\sqrt{N'_e}}{(\tilde{C}_H^{-1} \delta, \delta)^{1/2}} P_i Z^T \delta | \tilde{C}_H^{-1} \sim \mathcal{N}(0, P_i^2).$$

Denote $v = \frac{\sqrt{N'_e}}{(\tilde{C}_H^{-1} \delta, \delta)^{1/2}} P_i Z^T \delta$. Since $\text{rank } P_i < M - N_d$, the vector has a singular multivariate normal distribution. There exists a unique orthogonal matrix Q_i for each of the projectors P_i such that

$$Q_i P_i^2 Q_i^T = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix},$$

and $Q_i Q_i^T = I$ [17]. Applying this transform to the vector, we obtain:

$$Q_i v | \tilde{C}_H \sim \mathcal{N}(0, Q_i P_i^2 Q_i^T).$$

Denote the resulting covariance, $Q_i P_i^2 Q_i^T = L_i$. Note that L_i is a symmetric projection matrix, and $L_i = L_i L_i^T$.

By definition of the multivariate normal distribution, a normally distributed vector $u \sim \mathcal{N}(\mu, G)$ can be written as $u = \mu + R\zeta$, where $\zeta \sim \mathcal{N}(0, I)$ and $RR^T = G$. Since $Q_i P_i^2 Q_i^T = L_i L_i^T = L_i$, we write for the resulting vector, $Q_i v$:

$$Q_i v = Q_i P_i^2 Q_i^T \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_{M-N_d} \end{pmatrix} = \begin{pmatrix} \xi_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

where $\xi_k | \tilde{C}_H \sim \mathcal{N}(0, 1), k = 1, \dots, M - N_d$. The inner product (v, v) is then

$$\begin{aligned} (v, v) &= (Q_i^T Q_i v, v) = (Q_i v, Q_i v) \\ &= (\xi_1, 0, \dots, 0) \begin{pmatrix} \xi_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \xi_1^2. \end{aligned}$$

Since ξ_1 is marginally normal, the i th term of the sum has a marginal χ_1^2 distribution:

$$\frac{N'_e (P_i Z^T \delta, P_i Z^T \delta)}{(\tilde{C}_H^{-1} \delta, \delta)} | \tilde{C}_H \sim \chi_1^2. \quad (32)$$

After dividing by $\frac{(\tilde{C}_H^{-1} \delta, \delta)}{N'_e}$, the sum in Eq. 31 becomes

$$\begin{aligned} \frac{N'_e}{(\tilde{C}_H^{-1} \delta, \delta)} \sum_{i=1}^{M-N_d} \sigma_i^2 (P_i Z^T \delta, P_i Z^T \delta) \\ = \frac{N'_e}{(\tilde{C}_H^{-1} \delta, \delta)} \sum_{i=1}^{M-N_d} \sigma_i^2 \chi_1^2. \end{aligned}$$

We will call the weighted sum of the squared standard normal variables a generalized χ^2 distribution denoted as $\chi_{C_{2,2}}^2$, with corresponding density function $p_{\chi_{C_{2,2}}^2}$. Replacing the marginal distribution of the squared energy norm in Eq. 24 with the obtained distribution and using representation 22, we write for the conditional distribution, $P(\|\hat{\theta}^a - \theta^a\|^2 < \varepsilon | \tilde{C}_H)$:

$$\begin{aligned} P \left(\frac{(N'_e Z C_{2,2} Z^T \delta, \delta)}{(\tilde{C}_H^{-1} \delta, \delta)} < \varepsilon' \frac{(N'_e C_H^{-1} \delta, \delta)}{(\tilde{C}_H^{-1} \delta, \delta)} | \tilde{C}_H \right) \\ = P \left(\xi_{\chi_{C_{2,2}}^2} < \varepsilon' \frac{(N'_e C_H^{-1} \delta, \delta)}{(\tilde{C}_H^{-1} \delta, \delta)} | \tilde{C}_H \right) = \int_0^{\varepsilon' \frac{(N'_e C_H^{-1} \delta, \delta)}{(\tilde{C}_H^{-1} \delta, \delta)}} p_{\chi_{C_{2,2}}^2}(y) dy, \end{aligned}$$

where $\varepsilon' = \frac{\varepsilon}{(\tilde{C}_H^{-1} \delta, \delta)}$ and $\xi_{\chi_{C_{2,2}}^2}$ denotes the random value having a generalized χ^2 distribution.

Integrating over all \tilde{C}_H as in Eq. 26 and simplifying the resulting integral as in Eq. 27, we may now reformulate Theorem 1 for the regular Euclidean norm.

Theorem 2 Let $\bar{\theta}^a$ and θ^a be the outcomes from the analysis step of the EnKF and the Kalman filter respectively. Let $\delta = d - H\bar{\theta}^f$. Then

$$P(\|\bar{\theta}^a - \theta^a\|^2 < \varepsilon) = \int_0^\infty p_{\chi^2_{N_e - N_d + 1}}(z) \int_0^{\frac{z\varepsilon}{(C_H^{-1}\delta, \delta)}} p_{\chi^2_{C_{2,2}}}(y) dy dz. \quad (33)$$

By utilizing the distribution for the Euclidean norm of the sampling error, the behavior of the EnKF can be more easily assessed. Computing the double integral in Eq. 33 is a little more elaborate in comparison with Eq. 28 because of the generalized χ^2 distribution involved, but the computational effort is comparable. For reference on computation of the generalized χ^2 distribution, see, e.g., [5].

4 Numerical experiments

In this section, we demonstrate how Theorem 2 can be used to assess the reliability of EnKF. We assume that the result of the forecast stage is Gaussian in order to fulfill the requirements of Theorem 2. That is, we do not consider the additional difficulties the EnKF might encounter if the forecast distribution is not Gaussian.

We will investigate the effect of the numerical parameters included in the distribution 33, such as the system size, M , the number of ensemble members, N_e , and the number of observations, N_d . We will also investigate the effect of changing the structure of the observation operator, H .

For illustrative purposes, we assume the system state to be a $N \times N$ Gaussian random field, $N = 50$. We set an arbitrary reference state, θ (Fig. 1a). The reference state is used as a mean in the prior state distribution, $\mathcal{N}(\theta, C)$. The covariance matrix $C(N^2 \times N^2)$ is based on the spherical variogram model [28], with correlation range $r = 40$ and sill $s = 1$. The expression for the covariance matrix is the following:

$$C_{i,j} = \begin{cases} s \left(1 - \frac{3h}{2r} + \frac{h^3}{2r^3} \right), & \text{if } h < r, \\ 0, & \text{otherwise} \end{cases},$$

where

$$i = k_1 + N(l_1 - 1), k_1 = 1, \dots, N, l_1 = 1, \dots, N,$$

$$j = k_2 + N(l_2 - 1), k_2 = 1, \dots, N, l_2 = 1, \dots, N,$$

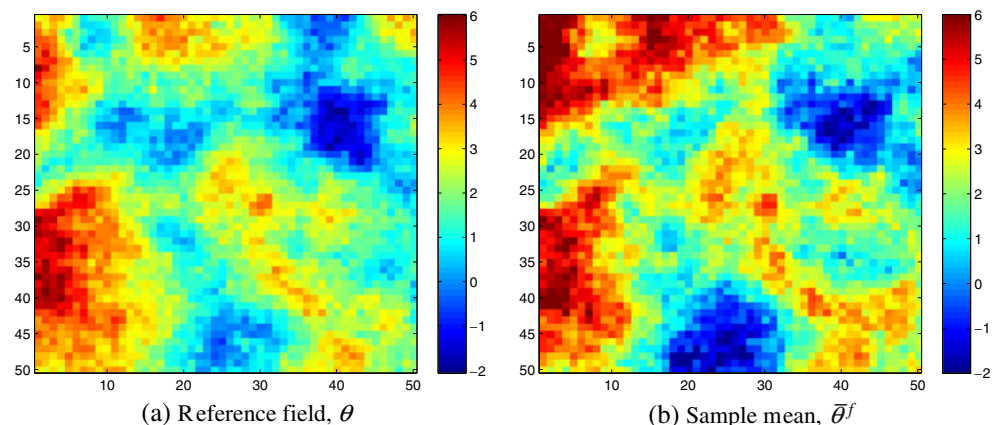
$$h^2 = (k_1 - k_2)^2 + (l_1 - l_2)^2.$$

From the latter distribution $\mathcal{N}(\theta, C)$, we draw N_e independent samples for each experiment we perform, denoting the sample mean as $\bar{\theta}^f$. The sample mean for $N_e = 100$ is shown in Fig. 1b.

In order to use the result of Theorem 2, we use the sample mean $\bar{\theta}^f$ as a prior for data assimilation for every experiment we perform. This allows us to concentrate on the error introduced at the analysis step. The sample mean is fixed for every experiment and is assumed to be a poorly known system state. We measure the reference field, θ , at certain locations and update the prior in order to improve the estimate.

We will consider different data assimilation scenarios in order to demonstrate the effect of each parameter characterizing the distribution of the sampling error norm (Eq. 33). For each considered case, we build the empirical distribution of the relative sampling

Fig. 1 Experiment setup. **a** Reference field, θ . **b** Sample mean, $\bar{\theta}^f$

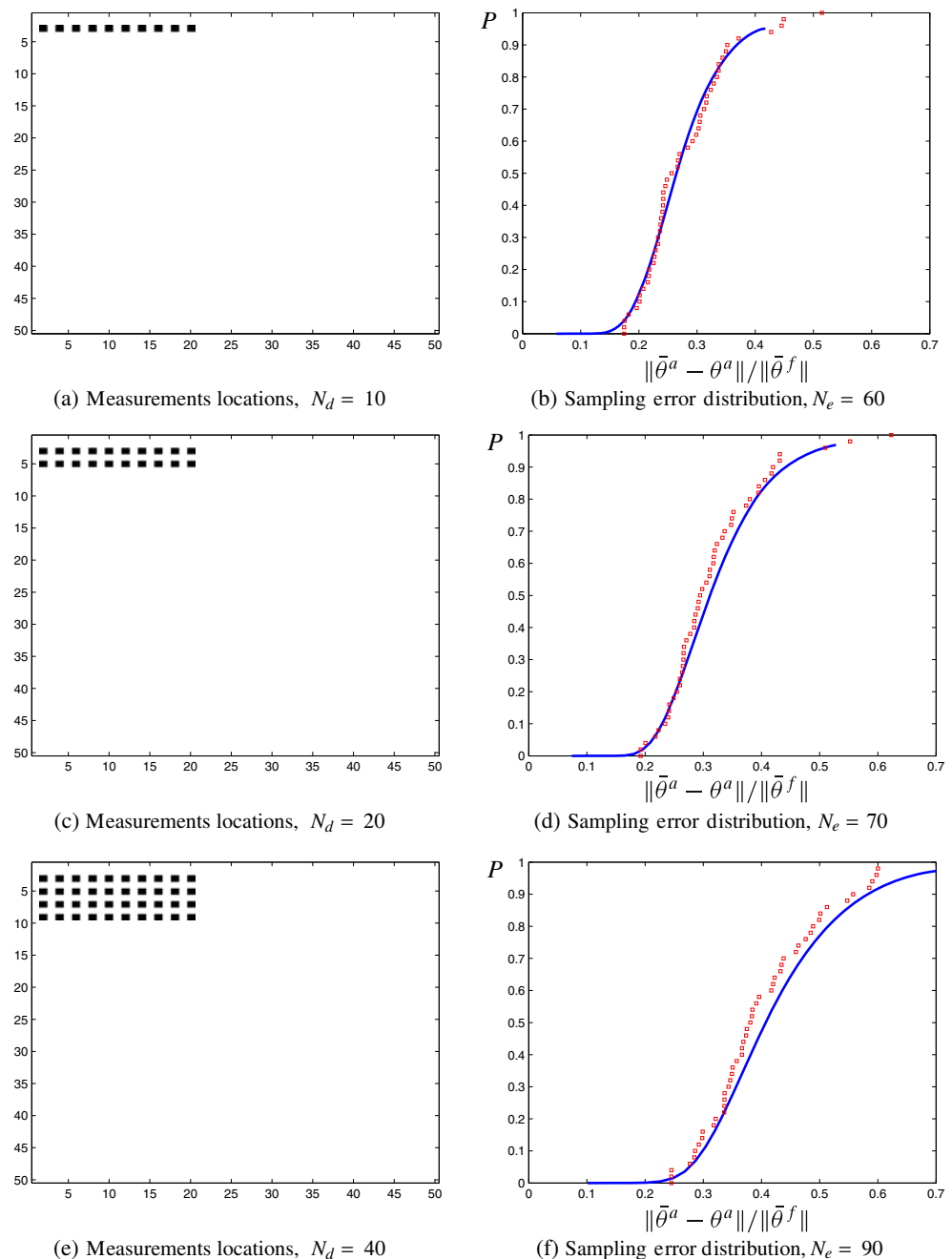


error, $\|\bar{\theta}^a - \theta^a\|/\|\bar{\theta}^f\|$. We use 50 points to build the empirical distribution. For this, we generate 50 independent realizations of the sample covariance matrix. Each realization is a sample drawn from the Wishart distribution, $W_M(N'_e, \frac{C}{N'_e})$. Each of the 50 matrices is then used in Eqs. 10–11, and the sampling error norm is calculated using Eq. 17, divided then by $\|\bar{\theta}^f\|$. Resulting values of the sampling error norm are then used to plot the empirical distribution. Even though 50 empirical realizations of the sampling error norm is relatively few, the number is sufficient for illustrative purposes, such

that empirical and analytical distributions can easily be distinguished.

The analytical expression for the sampling error given in Theorem 2 is compared with the empirical distribution. For a given measurement layout and corresponding matrix H , we calculate the quadratic form $(C_H^{-1}\delta, \delta)$ and use formula 33, thus building the cumulative distribution function for the squared sampling error $\|\bar{\theta}^a - \theta^a\|^2$. A simple variable transform in Eq. 33 gives the distribution for the relative sampling error, $\|\bar{\theta}^a - \theta^a\|/\|\bar{\theta}^f\|$.

Fig. 2 Distribution of the error norm, $N_e - N_d = 50$. **a** Measurements locations, $N_d = 10$. **b** Sampling error distribution, $N_e = 60$. **c** Measurements locations, $N_d = 20$. **d** Sampling error distribution, $N_e = 70$. **e** Measurements locations, $N_d = 40$. **f** Sampling error distribution, $N_e = 90$



4.1 Experiment 1

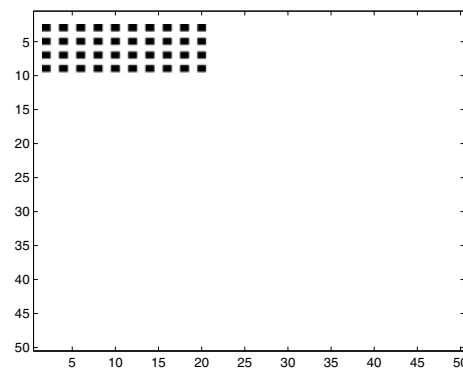
In this section, we consider the case when $N_e - N_d$ is constant and equal to 50. For this purpose, we vary $M - N_d$ by varying the number of measurements and, correspondingly, the ensemble size. We concentrate the measurement locations at the upper left corner of the field.

The figures on the left-hand side (Fig. 2a, c, e) illustrate the measurement locations, while the figures

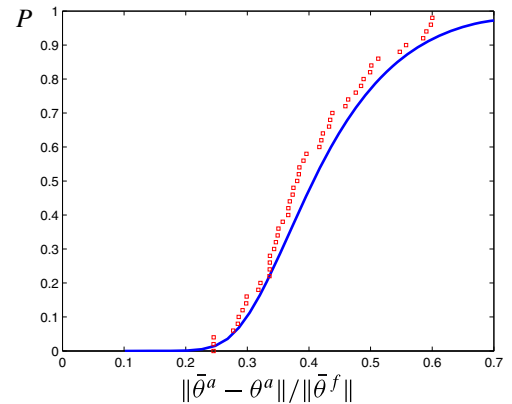
on the right-hand side (Fig. 2b, d, f) display the corresponding error distributions. The solid blue line corresponds to the cumulative distribution function of the relative sampling error, and the dashed red line is the empirical distribution for the same variable.

As seen from the resulting distributions, increasing the number of measurement locations leads to increase of the sampling error despite the increase of the ensemble size. The effect seen in this experiment arises due to changes in spectral properties of both C_H and $C_{2.2}$

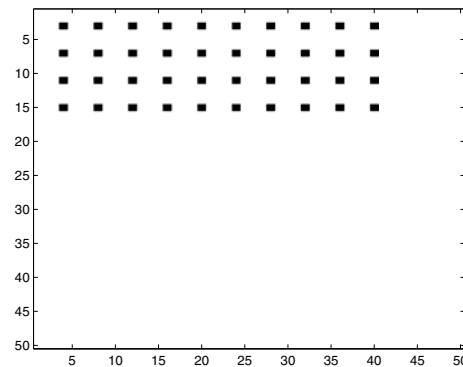
Fig. 3 Distribution of the error norm, $M - N_d = 2460$, $N_e - N_d = 50$, varied measurement locations. **a** Measurements locations, dense layout. **b** Sampling error distribution, dense layout. **c** Measurements locations, wide layout. **d** Sampling error distribution, wider layout. **e** Measurements locations, spread layout. **f** Sampling error distribution, spread layout



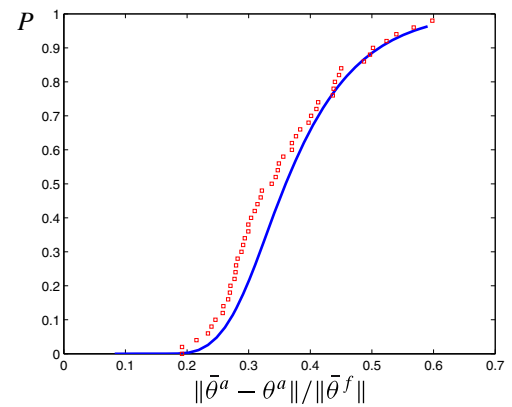
(a) Measurements locations, dense layout



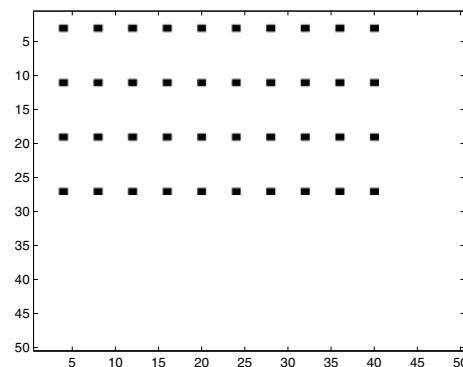
(b) Sampling error distribution, dense layout



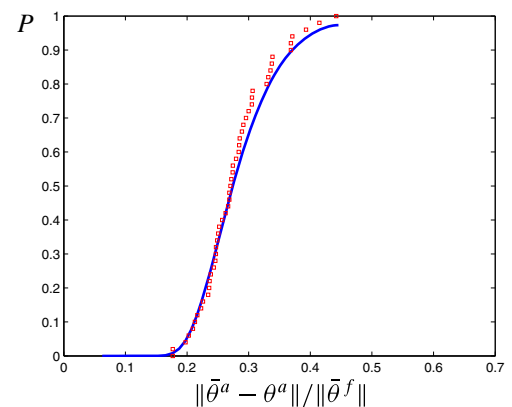
(c) Measurements locations, wide layout



(d) Sampling error distribution, wider layout



(e) Measurements locations, spread layout



(f) Sampling error distribution, spread layout

matrices, when increasing the number of the measurements, see Eq. 33.

4.2 Experiment 2

In this experiment, we consider a fixed ensemble size, $N_e = 90$, and a fixed number of measurements, $N_d = 40$. Since M is constant, $M - N_d$ is also constant. We vary the locations of the measurements and build the corresponding distributions.

As before, the left-hand side plots in Fig. 3 show the measurement layouts, and the plots on the opposite demonstrate the corresponding distributions. As seen from the figures, spreading the measurements across the field implies a smaller sampling error. Since we vary the structure of the measurement matrix, H , the matrices $C_{2,2}$ and C_H^{-1} vary correspondingly. The singular values of $C_{2,2}$ and the matrix C_H^{-1} are the only parameters affecting the sampling error norm distribution in this experiment. Thus, the spreading of the measurements improves both spectral properties of the $C_{2,2}$ matrix and the condition number of the matrix C_H , making the result of the EnKF analysis step a better estimate.

4.3 Experiment 3

The purpose of the next experiment is to hold $M - N_d$ constant, varying $N_e - N_d$ by changing only the ensemble size. We chose to consider the case with 40 measurements and dense measurement layout.

Figure 4a shows the measurement layout, and in Fig. 4b, the distributions corresponding to the different ensembles are displayed. As expected, increasing the ensemble size leads to a smaller sampling error norm

due to improved approximation of the true covariance matrix.

5 Summary and outlook

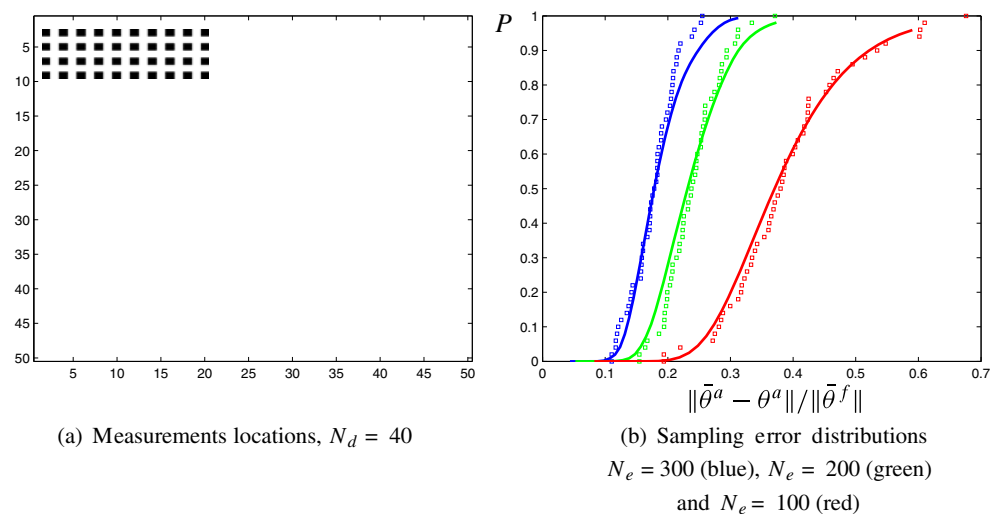
In Kovalenko et al., submitted, we derived the distribution for a particular problem-dependent energy norm of the sampling error. This energy norm was not easy to interpret, since the closeness of the EnKF solution to the reference was to a large degree determined by the spectral properties of the corresponding energy matrix.

In the current article, we derived the distribution for the Euclidean norm of the sampling error under the same assumptions of negligible errors and Gaussian forecast. The distribution depends on $N_e - N_d$, $M - N_d$, and covariance properties which can vary with these locations of the measurements.

We have demonstrated the use of the distribution on several numerical experiments, studying the effect of these parameters. Increasing the ensemble size and, correspondingly, $N_e - N_d$ reduces the sampling error norm. Increasing the number of observation, N_d , and keeping the difference, $N_e - N_d$, constant lead to a bigger sampling error. Spreading the observations across the measured field reduces the sampling error. Inclusion of a measurement error small enough that both $(C_H + C_d)^{-1} \approx C_H^{-1}$ and $(\tilde{C}_H + C_d)^{-1} \approx \tilde{C}_H^{-1}$ will allow one to apply the results of the current article.

The work on obtaining similar results in more realistic setting is currently being conducted. Taking into account measurement errors and obtaining a distribution for the EnKF deviation is of great interest. Another extension would be to study the EnKF error in the

Fig. 4 Distribution of the error norm, $M - N_d = 2,460$. **a** Measurements locations, $N_d = 40$. **b** Sampling error distributions $N_e = 300$ (blue), $N_e = 200$ (green), and $N_e = 100$ (red)



case of several time steps. Such results will significantly improve the understanding of the EnKF reliability in applications.

Acknowledgements The authors would like to acknowledge financial support to perform this study from the Research Council of Norway (Petromaks), Rocksource, Statoil, and Total.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

1. Aanonsen, S.I., Aavatsmark, I., Barkve, T., Cominelli, A., Gonard, R., Gosselin, O., Kolasinski, M., Reme, H.: Effect of scale dependent data correlations in an integrated history matching loop combining production data and 4D seismic data. In: Proc. SPE Reservoir Simulation Symposium, paper SPE 79665. Houston, Texas (2003)
2. Aanonsen, S.I., Nævdal, G., Oliver, D.S., Reynolds, A.C., Valles, B.: The ensemble Kalman filter in reservoir engineering—a review. *SPE J.* **14**(3), 393–412 (2009)
3. Bennett, A.F.: Inverse Modeling of the Ocean and Atmosphere. Cambridge University Press, Cambridge (2002)
4. Chen, C., Wang, Y., Li, G., Reynolds, A.: Closed-loop reservoir management on the Brugge test case. *Computational Geosciences* **14**, 691–703 (2010)
5. Davies, R.B.: Algorithm as 155: the distribution of a linear combination of χ^2 random variables. *J. R. Stat. Soc., Ser. C, Appl. Stat.* **29**(3), 323–333 (1980)
6. Dong, Y., Oliver, D.S.: Quantitative use of 4D seismic data for reservoir description. *SPE J.* **10**(1), 91–99 (2005)
7. Evensen, G.: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.* **99**, 10143–10162 (1994)
8. Evensen, G.: The ensemble Kalman filter: theoretical formulation and practical implementation. *Ocean Dyn.* **53**, 343–367 (2003)
9. Evensen, G.: Sampling strategies and square root analysis schemes for the EnKF. *Ocean Dyn.* **54**, 539–560 (2004)
10. Evensen, G.: Data Assimilation: The Ensemble Kalman Filter, 2nd edn. Springer, New York (2009)
11. Evensen, G., van Leeuwen, P.J.: An ensemble Kalman smoother for nonlinear dynamics. *Mon. Weather Rev.* **128**(6), 1852–1867 (2000)
12. Fahimuddin, A., Mannseth, T., Aanonsen, S.I.: Effect of large number of measurements on the performance of EnKF model updating. In: Proceedings of the 11th European Conference on the Mathematics of Oil Recovery (ECMOR XI). Bergen, Norway (2008)
13. Feng, T., Mannseth, T.: Impact of time-lapse seismic data for permeability estimation. *Comput. Geosci.* **14**, 705–719 (2010)
14. Furrer, R., Bengtsson, T.: Estimation of high-dimensional prior and posterior covariance matrices in Kalman filter variants. *J. Multivar. Anal.* **98**(2), 227–255 (2007)
15. Gupta, A.K., Nagar, D.K.: Matrix variate distributions. Chapman & Hall/CRC Monographs and Surveys in Pure and Applied Mathematics, vol. 104. Chapman & Hall/CRC, Boca Raton (2000)
16. Hayakawa, T.: On the distribution of a quadratic form in a multivariate normal sample. *Ann. Inst. Stat. Math.* **18**, 191–201 (1966)
17. Horn, R.A., Johnson, C.R.: Matrix Analysis. Cambridge University Press, Cambridge (1990)
18. Houdekamer, P.L., Mitchell, H.L.: Data assimilation using an ensemble Kalman filter technique. *Mon. Weather Rev.* **126**(3), 796–811 (1998)
19. Kalman, R.E.: A new approach to linear filtering and prediction problems. *Trans. Am. Soc. Mech. Eng.—J. Basic Eng.* **82**(Series D), 35–45 (1960)
20. Kepert, J.D.: On ensemble representation of the observation-error covariance in the ensemble Kalman filter. *Ocean Dyn.* **54**, 561–569 (2004)
21. Kovalenko, A., Mannseth, T., Nævdal, G.: Error estimate for the ensemble Kalman filter update step. In: Proceedings of the 12th European Conference on the Mathematics of Oil Recovery (ECMOR XII). Oxford, UK (2010)
22. Nævdal, G., Johnsen, L.M., Aanonsen, S.I., Vefring, E.H.: Reservoir monitoring and continuous model updating using ensemble Kalman filter. *SPE J.* **10**(1), 66–74 (2005)
23. Nævdal, G., Mannseth, T., Vefring, E.H.: Near-well reservoir monitoring through ensemble Kalman filter. In: The 2002 SPE/DOE Improved Oil Recovery Symposium, Paper SPE 75235. Tulsa, Oklahoma (2002)
24. Oliver, D., Chen, Y.: Ensemble-based closed-loop optimization applied to Brugge field. *SPE Reserv. Evalu. Eng.* **13**(1), 56–71 (2010)
25. Oliver, D., Chen, Y.: Recent progress on reservoir history matching: a review. *Comput. Geosci.* **15**, 1–37 (2010)
26. Skjervheim, J.A., Evensen, G., Aanonsen, S.I., Ruud, B.O., Johansen, T.A.: Incorporating 4D seismic data in reservoir simulation models using ensemble Kalman filter. *SPE J.* **12**(3), 282–292 (2007)
27. Srivastava, M.S.: Singular Wishart and multivariate beta distributions. *Ann. Stat.* **31**(5), 1537–1560 (2003)
28. Wackernagel, H., Bertino, L.: Geostatistics and Sequential Data Assimilation. Springer, New York (2005)